



AI Based Air Pollution Monitoring and Prediction System

Ms.K. Madhumitha¹, M Mohana Priya², R Priya³, D Namitha⁴

¹Assistant Professor, Department of Computer Science and Engineering
Kongunadu College of Engineering and Technology Tamilnadu,India

^{2,3,4}Department of Computer Science and Engineering Kongunadu College of Engineering and
Technology Tamilnadu,India

Abstract. Air pollution is one of the most serious environmental and health concerns in modern urban societies. This study proposes an AI-driven Air Pollution Monitoring and Prediction System that enables continuous observation and accurate forecasting of air quality levels. The system employs IoT-based sensors to gather real-time data on pollutants such as PM2.5, PM10, CO, NO₂, and SO₂, along with meteorological parameters including temperature and humidity. The collected data is transmitted to a cloud environment for preprocessing and analysis. Advanced machine learning and deep learning techniques, particularly time-series models like LSTM and regression algorithms, are utilized to predict future Air Quality Index (AQI) levels. The system also incorporates interactive dashboards and automated alert mechanisms to inform authorities and the public about potential pollution risks. Performance evaluation indicates higher prediction accuracy and lower error margins compared to conventional statistical approaches. The proposed framework supports data-driven environmental management, timely intervention strategies, and sustainable urban development while promoting improved public health outcomes.

Keywords: Air Quality Monitoring, Artificial Intelligence, IoT, Machine Learning, Deep Learning, AQI Prediction, LSTM, Environmental Monitoring, Smart Cities.

I. Introduction

Air pollution has emerged as a critical environmental and public health issue worldwide, driven by rapid urbanization, industrial growth, increasing vehicular traffic, and excessive energy consumption. Pollutants such as particulate matter (PM2.5 and PM10), carbon monoxide (CO), nitrogen dioxide (NO₂), and sulfur dioxide (SO₂) accumulate in the atmosphere and adversely affect air quality. Prolonged exposure to these contaminants contributes to respiratory illnesses, cardiovascular complications, reduced productivity, and environmental imbalance. As urban populations continue to expand, ensuring effective air quality management has become an essential responsibility for governments and environmental authorities.

Conventional air quality monitoring approaches primarily depend on fixed monitoring stations and traditional statistical analysis methods. While these systems provide standardized measurements, they often involve high infrastructure costs and limited geographical coverage. Additionally, traditional forecasting techniques may struggle to model the complex and nonlinear relationships among pollutants and meteorological variables, resulting in less accurate predictions. These limitations highlight the need for advanced, adaptive, and real-time monitoring solutions.



The emergence of Artificial Intelligence (AI), Machine Learning (ML), and the Internet of Things (IoT) has significantly transformed environmental monitoring practices. IoT-enabled sensors facilitate continuous data acquisition of pollutant concentrations and atmospheric parameters such as temperature and humidity. The collected data is transmitted to centralized cloud platforms for storage, processing, and analysis. AI-based models are capable of learning intricate patterns from large datasets, identifying trends, and generating reliable predictions. Deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, are well suited for analyzing time-series data and capturing temporal dependencies in air pollution trends.

An AI-Based Air Pollution Monitoring and Prediction System combines sensing technologies with intelligent data analytics to create an integrated and efficient framework. Such systems provide real-time Air Quality Index (AQI) updates, predictive insights, and early warning notifications. Through visualization dashboards and automated alerts, stakeholders can take timely preventive measures to mitigate pollution risks.

This study aims to develop a scalable and accurate AI-driven system that enhances forecasting performance and supports proactive environmental management. By integrating advanced computational intelligence with smart sensing infrastructure, the proposed approach contributes to sustainable urban development and improved public health protection.

II. Related Works

Artificial Intelligence-Based Monitoring and Forecasting of Urban Air Pollution in Smart Cities investigates AI-driven methods for urban air pollution tracking and forecasting using data collected from city-wide sensor networks. Four machine learning algorithms — Decision Tree, Random Forest, Support Vector Machine, and Artificial Neural Network — were evaluated for predicting levels of key pollutants such as PM_{2.5}, PM₁₀, NO₂, and CO. Experimental outcomes show that ANN models achieved the best prediction accuracy (~94.8%), followed by Random Forest and SVM, demonstrating the effectiveness of AI in modeling complex pollutant behavior. The work emphasizes the advantages of automated forecasting tools over traditional statistical approaches, particularly in smart city environments where rapid insights and adaptive responses are required for environmental management and public health protection.

Integrating AI With IoT Sensors for Real-Time Air Quality Monitoring and Prediction Supriya Kumari proposes a real-time air quality monitoring and forecasting system by integrating IoT sensor networks with AI models. A deployment of 50 IoT sensors across North Delhi continuously captured PM_{2.5}, PM₁₀, NO₂, and CO data at two-minute intervals. The collected data streams were processed using machine learning algorithms, enabling high-resolution spatial and temporal air quality predictions. Results demonstrated that the AI-IoT framework can produce accurate, cost-effective short-term forecasts that outperform many conventional setups, offering a scalable solution for urban air pollution management and health advisories.

FuXi-Air: Urban Air Quality Forecasting Based on Emission-Meteorology-Pollutant Multimodal Machine Learning ,Hao Li.et.al.,introduces a novel multimodal forecasting model that combines emission inventories, meteorological forecasts, and pollutant



measurements for high-precision urban air quality forecasting. The model uses an autoregressive framework along with frame interpolation to predict six major pollutants at an hourly resolution up to 72 hours ahead, achieving superior efficiency and accuracy compared to traditional numerical models. Multimodal data fusion is shown to improve forecasting performance under varying pollution conditions, making it suitable for smart city air quality risk management.

PlumeNet: Large-Scale Air Quality Forecasting Using a Convolutional LSTM Network PlumeNet focuses on large-scale forecasting of key pollutants such as NO₂, O₃, PM_{2.5}, and PM₁₀ using a convolutional LSTM architecture. The model incorporates weather forecasts, existing monitoring station data, and established physical model outputs. By applying convolutional LSTM layers, PlumeNet achieves reliable long-term air quality forecasts with reduced computational requirements compared to standard physics-based models. The study emphasizes scalability and frequent updates based on new measurements, enabling dynamic air quality predictions crucial for public goods like health risk warnings and environmental planning.

VayuAnukulani: Adaptive Memory Networks for Air Pollution Forecasting Madaan,et.al., presents VayuAnukulani, an adaptive end-to-end forecasting system for predicting levels of pollutants in Delhi using bidirectional LSTM with attention mechanisms. Trained on historical and real-time data, the model predicts concentrations for major pollutants like NO₂, PM_{2.5}, and PM₁₀ up to 24 hours ahead. Results show that the adaptive memory network outperforms traditional baseline models, providing enhanced prediction accuracy useful for rapid response and public health advisories in highly polluted urban environments.

AirNet: Predictive Machine Learning Model for Air Quality Forecasting Using Web Interface uses multiple traditional machine learning classifiers including Random Forest, Logistic Regression, Decision Tree, Naive Bayes, SVM, and K-Nearest Neighbors to predict concentrations of PM_{2.5}, O₃, CO, and NO₂. Designed as a web-integrated forecasting and alert system, it handles data preprocessing, outlier management, and generates AQI predictions. The study highlights the utility of ensemble and classical algorithms in supporting decision-making for urban air quality mitigation strategies, emphasizing practical deployment considerations and real-time alerting functionality.

Monitoring and Predicting Air Quality with IoT Devices Razvan Bogdan,et.al., integrates IoT sensors with machine learning models for real-time air quality index (AQI) prediction. Multiple environmental parameters (temperature, humidity, PM_{2.5}, PM₁₀) were collected and transmitted to cloud storage for preprocessing. Machine learning regression models implemented in TensorFlow provided AQI forecasts with improved responsiveness and accuracy. The study compares neural network and Random Forest approaches, demonstrating the feasibility of combined IoT and ML systems as practical tools for environmental monitoring and timely public health support.

Temporal Trends and Predictive Modeling of Air Pollutants in Delhi Ali Alsuwaiyan, et al. examines temporal trends and compares multiple AI forecasting methods for major air pollutants in Delhi. Using models such as regression and machine learning algorithms, it forecasts CO, NO, NO₂, O₃, SO₂, PM_{2.5}, PM₁₀, and NH₃ levels, analyzing model performance under different atmospheric conditions. The study provides insights



on pollutant dynamics and the relative effectiveness of AI models in addressing urban air quality challenges, offering a data-driven basis for environmental policy recommendations.

Improved Prediction Model for Daily PM_{2.5} Concentrations with Particle Swarm Optimization and BP Neural Network Umesh Kumar Lilhore, et.al., study proposes an enhanced prediction model for daily PM_{2.5} levels using a Backpropagation (BP) neural network optimized with Particle Swarm Optimization (PSO). By tuning network weights and thresholds optimally, the model achieves stable and accurate PM_{2.5} forecasts using seven key input features related to pollutant and meteorological variables. The integration of PSO improves convergence and prediction reliability compared to standalone neural network models.

Applications of Machine Learning and IoT for Outdoor Air Pollution Monitoring and Prediction: A Systematic Literature Review Ihsane Gryech, et.al., systematic review synthesizes research on combining machine learning and IoT for outdoor air pollution monitoring. It classifies prediction approaches into time-series, feature-based, and spatio-temporal methods and highlights challenges such as limited data diversity and context-specific features. The paper identifies gaps in coverage and proposes future research directions, emphasizing how hybrid IoT-ML systems can enhance accuracy, lower costs, and improve urban air quality management globally.

III. Proposed Method

The proposed AI-Based Air Pollution Monitoring and Prediction System aims to deliver continuous air quality surveillance, accurate forecasting, and timely alert generation through an integrated intelligent framework. The system combines IoT sensor networks, cloud-based data management, and advanced machine learning techniques to create a reliable and scalable environmental monitoring solution.

In the data collection layer, IoT-enabled sensors are strategically installed across different locations to measure major air pollutants such as PM_{2.5}, PM₁₀, CO, NO₂, and SO₂, along with atmospheric parameters like temperature and humidity. These sensors transmit real-time readings to a centralized cloud platform using wireless communication technologies. To ensure data reliability, preprocessing operations including noise removal, missing data handling, normalization, and anomaly detection are performed before analysis.

The processed data is stored in a cloud database for efficient access and large-scale analytics. In the prediction stage, the system utilizes both traditional machine learning models and deep learning approaches. Algorithms such as Random Forest Regression and Support Vector Regression provide baseline estimations, while Long Short-Term Memory (LSTM) networks are employed to capture temporal patterns in time-series air quality data. The LSTM model analyzes historical pollutant levels and meteorological factors to generate short-term and medium-term AQI forecasts.

Feature engineering and correlation analysis are conducted to identify the most impactful environmental variables, improving model performance. Hyperparameter optimization

tion techniques further enhance prediction accuracy and minimize overfitting. The system is designed to update and retrain models periodically using newly collected data, ensuring adaptability to seasonal fluctuations and evolving pollution trends.

At the application layer, an interactive web and mobile dashboard presents real-time AQI readings, pollutant concentration trends, historical comparisons, and predictive insights. Visual indicators such as color-coded AQI levels enable quick interpretation. Additionally, automated alerts via SMS or email are triggered when pollutant concentrations exceed predefined safety limits. Overall, the proposed system offers improved accuracy, scalability, and responsiveness compared to conventional monitoring methods. By integrating IoT infrastructure with AI-driven predictive analytics, it supports data-informed environmental management, proactive pollution control, and enhanced public health safety.

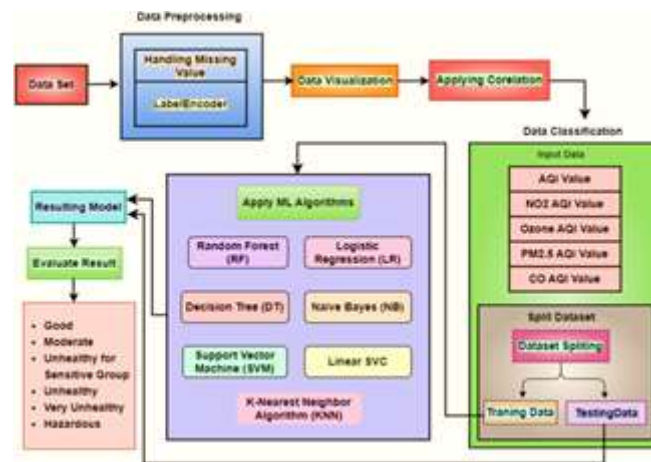


Fig.1.System Architecture

The AI-Based Air Pollution Monitoring and Prediction System is organized into several interconnected modules to ensure smooth data flow, intelligent analysis, and effective user communication.

1. Environmental Data Collection Module:

This module is designed to capture real-time air quality data through IoT-enabled sensors deployed across various locations. The sensors continuously measure pollutants such as PM2.5, PM10, CO, NO₂, and SO₂, along with environmental factors like temperature and humidity. Data is transmitted periodically to a centralized system using wireless communication technologies. This module guarantees continuous observation and broad geographic monitoring coverage.

2. Data Processing Module:

Since raw sensor readings may contain noise, incomplete entries, or irregular values, this module performs preprocessing tasks including data cleaning, anomaly detection, normalization, and missing value handling. It converts unstructured sensor outputs into a structured dataset suitable for machine learning analysis. By enhancing data quality, this module improves the robustness of predictive modeling.

3. Cloud Storage and Data Management Module:

Processed data is securely stored in a cloud-based database within this module. It ensures scalable storage capacity, secure data handling, and efficient retrieval of historical records. Proper indexing and management techniques enable fast access to large volumes of environmental data for training and analysis purposes.

4. Prediction and Analytical Module:

This module forms the core intelligence of the system. It employs machine learning algorithms such as Random Forest and Support Vector Regression, along with deep learning models like Long Short-Term Memory (LSTM) networks, to analyze time-series pollution data. By identifying trends, seasonal patterns, and variable relationships, the module forecasts future Air Quality Index (AQI) levels. Periodic model updates allow the system to adapt to changing environmental conditions.

5. Alert Generation Module:

This component automatically triggers warnings when pollutant levels exceed predefined safety thresholds. Notifications are delivered via SMS, email, or mobile alerts to inform users and regulatory authorities, enabling quick response and mitigation measures.

6. Visualization and User Interaction Module:

The final module provides an interactive dashboard that displays live AQI readings, pollutant concentration charts, historical comparisons, and predicted outcomes. Color-coded AQI categories help users interpret air quality conditions easily.3. Communication and Cloud Integration Module:

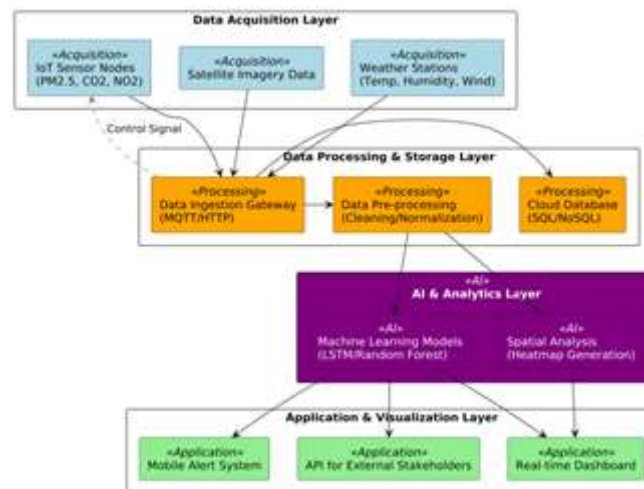


Fig.2. Methodology workflow of the Intelligent Air Pollution Detection and Prediction Framework

Overall Working Flow of the Proposed System:

The overall workflow of the AI-Based Air Pollution Monitoring and Prediction System begins with real-time data collection through IoT-enabled air quality sensors deployed across multiple locations. These sensors continuously measure pollutant concentrations



such as PM_{2.5}, PM₁₀, CO, NO₂, and SO₂, along with meteorological parameters including temperature and humidity. The collected environmental data is transmitted at regular intervals through wireless communication networks to a centralized cloud platform for further processing.

Once the data reaches the cloud server, it undergoes preprocessing to improve quality and consistency. This includes removing noise, handling missing values, detecting and eliminating outliers, and normalizing the dataset. The cleaned and structured data is then stored in a secure cloud database, where historical records are maintained for long-term analysis and model training. Proper data management ensures scalability and efficient retrieval of large datasets.

After preprocessing and storage, the analytical phase begins. Machine learning and deep learning algorithms are applied to identify patterns, correlations, and temporal trends in air pollution data. Regression-based models provide baseline predictions, while advanced time-series models such as Long Short-Term Memory (LSTM) networks capture sequential dependencies and seasonal variations. The trained models analyze historical pollutant behavior along with meteorological influences to forecast future Air Quality Index (AQI) values for short-term and medium-term intervals.

The system continuously evaluates prediction performance using accuracy metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Based on evaluation results, hyperparameters are tuned and models are periodically retrained using newly collected data to maintain adaptability under changing environmental conditions. The predicted AQI values and real-time pollutant levels are then displayed through an interactive web or mobile dashboard. Visual representations such as graphs, charts, and color-coded AQI categories allow users to easily interpret air quality conditions. If the forecasted or real-time pollution levels exceed predefined safety thresholds, the system automatically generates alerts via SMS, email, or mobile notifications.

Through this integrated workflow, the system ensures continuous monitoring, intelligent forecasting, timely alert generation, and informed decision-making, thereby supporting proactive air pollution management and public health protection.

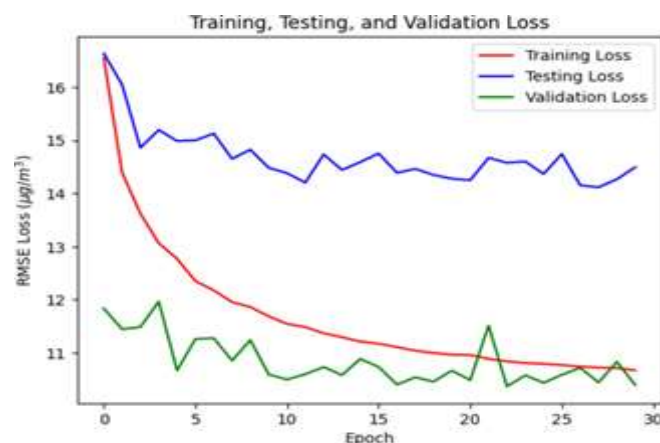


Fig.3. Performance Evaluation of Intelligent Air Pollution Detection and Prediction Framework



$$AQI = \frac{(I_{high} - I_{low})}{(C_{high} - C_{low})} \times (C - C_{low}) + I_{low}$$

The Air Quality Index (AQI) formula is used to convert raw pollutant concentration values into a standardized scale that indicates the level of health risk. Since different pollutants have different concentration ranges and health impacts, the AQI formula normalizes pollutant measurements into a uniform index scale, typically ranging from 0 to 500. The formula performs linear interpolation between defined breakpoint ranges. Each pollutant, such as PM_{2.5} or NO₂, has predetermined concentration thresholds established by environmental agencies.

Based on where the measured concentration falls within these breakpoints, the AQI value is calculated. The highest AQI among all pollutants is considered the overall AQI for a given time period. This standardized index helps authorities and the public easily interpret pollution levels as categories such as Good, Moderate, Unhealthy, or Hazardous. In the proposed AI system, AQI serves as the target variable for prediction models, enabling the system to forecast health risk levels rather than raw pollutant values alone.

References

1. M. T. Huque, J. J. Godhuli, S. M. R. Rahman Pushon, and E. Haque, "Internet of Things (IoT)-based smart irrigation system for sustainable agriculture," *J. Infor. Electr. Electron. Eng.*, vol. 4, no. 3, pp. 1–7, Nov. 2023.
2. Tran A. Khoa, M. M. Man, T.-Y. Nguyen, V. D. Nguyen, and N. H. Nam, "Smart agriculture using IoT multi-sensors: A novel watering management system," *J. Sens. Actuator Netw.*, vol. 8, no. 3, p. 45, Aug. 2019.
3. "The Smart IoT based Automated Irrigation System using Arduino UNO and Soil Moisture Sensor," *IEEE Xplore Conf. Proc.*, 2025.
4. "Smart Farming: Design and Implementation of an IoT-Based Automated Irrigation System for Precision Agriculture," *IEEE Conf. Proc.*, 2025.
5. "Efficient Smart Irrigation System using Internet of Things (IoT) for Agriculture Crops," *IEEE Conf. Proc.*, 2025.
6. "Revolutionizing Farming with IoT: Smart Irrigation System for Sustainable Agriculture," *IEEE Conf. Proc.*, 2025.
7. "Smart agriculture using IoT for automated irrigation, water and energy efficiency," *AerTech*, 2025.
8. M. V. Gurao and U. B. Vaidya, "IoT based Smart Irrigation System," *Int. J. Eng. Res. Technol.*, vol. 11, no. 11, Nov. 2022.
9. Dr. M. Kumari and A. K. Sah, "IoT enabled smart irrigation system, monitoring and water harvesting in different soils," *Int. J. Eng. Res. Technol.*, vol. 10, no. 03, Mar. 2021.
10. Prabjyot S. Solar and Manoj Mishra, "Smart irrigation system using IoT," *Int. J. Adv. Res., Ideas Innov. Technol.*, vol. 11, no. 5, Nov. 2025.
11. Kamogelo Taueatsoala et al., "TinyML-Enabled IoT for Sustainable Precision Irrigation," *arXiv*, Jan. 19, 2026.
12. T. Ahmadi Pargo, M. A. Shirazi, and D. Fadai, "Smart and Efficient IoT-Based Irrigation System Design: Utilizing a Hybrid Agent-Based and System Dynamics Approach," *arXiv*, Feb. 25, 2025.



13. A. Hasib and A. S. M. A. Sarkar Akib, "An IoT-Based Smart Plant Monitoring and Irrigation System with Real-Time Environmental Sensing, Automated Alerts, and Cloud Analytics," arXiv, Jan. 22, 2026.
14. Y. Emre Kunt, "Development of a Smart Autonomous Irrigation System Using IoT and AI," arXiv, Jun. 13, 2025.
15. X. Liu, Z. Zhao, and A. Rezaeipannah, "Intelligent and automatic irrigation system based on Internet of Things using fuzzy control technology," *Sci. Rep.*, vol. 15, Art. no. 14577, 2025.
16. A. Morchid, R. Jebabra, H. M. Khalid, R. E. Alami, H. Qjidaa, and M. O. Jamil, "IoT-based smart irrigation management system to enhance agricultural water security using embedded systems, telemetry data, and cloud computing," *Results Eng.*, 2024.
17. S. Patel, A. Sharma, and B. Singh, "Automated irrigation system using soil moisture sensors and Raspberry Pi," *Int. J. Automat. Agric.*, vol. 3, no. 2, pp. 55–62, 2022.
18. H. Kumar, L. Desai, and S. Verma, "IoT-Based Smart Irrigation System Using Soil Moisture Sensor and Raspberry Pi," *IEEE Internet of Things J.*, vol. 9, no. 3, pp. 1008–1015, 2023.
19. A. Roy, P. Bhattacharya, and R. Sen, "Design and implementation of an IoT-enabled automated irrigation system using soil moisture sensing," *IEEE Access*, vol. 10, pp. 2354–2362, 2022.
20. M. Agarwal, K. Reddy, and T. Singh, "Smart farming: An automated irrigation system using Raspberry Pi and IoT," *Agric. Eng. Today*, vol. 45, no. 4, pp. 32–38, 2023.
21. J. Singh, V. Gupta, and M. Patel, "An IoT-driven approach to automated irrigation using soil moisture sensors," *Int. J. Agric. Sci. Technol.*, vol. 8, no. 2, pp. 1–10, 2023.
22. "Automated Smart Irrigation System for Sustainable Water Management," *IEEE Conf. Proc.*, 2024.
23. "IoT in Precision Agriculture: Sensor Networks and Decision Algorithms," *IEEE Trans. Ind. Inf.*, 2023.
24. R. Zhang and Y. Liu, "Real-time IoT-enabled smart irrigation systems in precision agriculture," *Comput. Electron. Agric.*, vol. 173, p. 105423, 2020.
25. S. Borkar and R. Pandit, "Artificial intelligence and IoT for sustainable agriculture," *IEEE Trans. Ind. Inf.*, vol. 17, no. 8, pp. 2561–2571, 2021.
26. C. Bouras, A. Kollia, and A. Papazois, "SDN & NFV in 5G: Advancements and challenges for smart agriculture," *IEEE Innov. Clouds Internet Netw.*, pp. 107–111, 2017.
27. P. Kumar and D. Sharma, "AI-driven smart irrigation: A sustainable approach for water conservation," *IEEE Xplore*, 2021.
28. C. Rao, "Internet of Things in agriculture: Smart farming techniques," Springer, 2021.
29. H. Wang, "AI for precision agriculture: Smart technologies for sustainable farming," Wiley, 2020.
30. R. Smith and R. Patel, "IoT-based smart irrigation system for precision farming," *Int. J. Agric. Innov.*, vol. 10, no. 3, pp. 45–56, 2022.
31. S. Rahman and M. Jamil, "Smart irrigation techniques in modern agriculture: A review," *J. Agric. Eng.*, vol. 56, no. 2, pp. 81–90, 2020.



32. V. Gupta and M. Reddy, "Implementation of an IoT and machine learning-based smart irrigation system," *Smart Agric. J.*, vol. 7, no. 4, pp. 210–220, 2020.
33. C. Lee and J. Cho, "Smart water management for precision farming using IoT and AI," *J. Environ. Manage.*, vol. 281, p. 111963, 2021.
34. S. Singh and P. Sharma, "Integration of AI and IoT for smart agriculture," Springer, pp. 30–45, 2021.
35. P. Bhattacharya and R. Dey, "Advances in IoT-based irrigation systems: A comprehensive review," *IEEE Access*, vol. 9, pp. 12345–12362, 2021.